

SCAN, sequence search for SEA ID NO 12

Worley 10/751235

Page 1

=> fil reg; d que l4

FILE 'REGISTRY' ENTERED AT 15:12:09 ON 09 MAR 2006
USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.
PLEASE SEE "HELP USAGETERMS" FOR DETAILS.
COPYRIGHT (C) 2006 American Chemical Society (ACS)

Property values tagged with IC are from the ZIC/VINITI data file
provided by InfoChem.

STRUCTURE FILE UPDATES: 8 MAR 2006 HIGHEST RN 876273-86-8
DICTIONARY FILE UPDATES: 8 MAR 2006 HIGHEST RN 876273-86-8

New CAS Information Use Policies, enter HELP USAGETERMS for details.

TSCA INFORMATION NOW CURRENT THROUGH January 6, 2006

Please note that search-term pricing does apply when
conducting SmartSELECT searches.

*
* The CA roles and document type information have been removed from *
* the IDE default display format and the ED field has been added, *
* effective March 20, 2005. A new display format, IDERL, is now *
* available and contains the CA role and document type information. *
*

Structure search iteration limits have been increased. See HELP SLIMITS
for details.

REGISTRY includes numerically searchable data for experimental and
predicted properties as well as tags indicating availability of
experimental property data in the original document. For information
on property searching in REGISTRY, refer to:

<http://www.cas.org/ONLINE/UG/regprops.html>

L1 14471 SEA FILE=REGISTRY ABB=ON [AG]G.[DE]T[TS]/SQSP
L3 6 SEA FILE=REGISTRY ABB=ON LOPYAEDGSAVNMEAKFSQMTLDVIGLSLFN|VYTAL
KEAELRSTDLLPYWKIDALCKIVPRO|VIGLSLFNFDSLTDDSPVIEAVYTALKEA/SQSP
L4 5 SEA FILE=REGISTRY ABB=ON L1 AND L3

• = any amino acid
Seq ID 12
} overlapping 31aa fragments of Seq ID 1

=> d:seqid: l4: 1=59

(31aa / 77aa = 40%)

L4 ANSWER 1 OF 5 REGISTRY COPYRIGHT 2006 ACS on STN
RN 857687=96=8 REGISTRY *Use Registry # to match sequence to reference (printed*
CN Cytochrome P 450 (Arabidopsis thaliana gene LUT-1 isoenzyme CYP97C1 *beginning on*
precursor) (9CI) (CA INDEX NAME) *pg 5)*
OTHER NAMES:
CN 4: PN: US20050150002 SEQID: 4 claimed protein
FS PROTEIN SEQUENCE
SQL 539

SEQ 1 MESSLFSPSS SSYSSLFYAK PTRLLSPKPK FTFSIRSSIE KPKPKLETNS
51 SKSQSWVSPD WLTTLTRTSL SGKNDESGIP IANAKLDDVA DLLGGALFLP
101 LYKWMNEYGP IYRLAAGPRN FVIVSDPAIA KHVLRNYPKY AKGLVAEVSE

151 FLFGSGFAIA EGPLWTARRR AVVPSLHRRY LSVIVERVFC KCAERLVEKL
 201 ~~QPYAEDGSAV NMEAKFSQMT LDVIGLSLFN YNFDSLTTDS PVIEAVYTA~~
 251 ~~KEAELRSTD LPYWKIDALC KIVPROVKAE KAVTLIRETV EDLIAKCKEI~~
 301 VEREGERIND EEYVNDADPS ILRFLLASRE EVSSVQLRDD LLSMLVAGHE
 351 TTGSVLTWTL YLLSKNSSAL RKAQEEVDRV LEGRNPAFED IKELKYITRC
 401 INESMRLYPH PPVLIRRAQV PDILPGNYKV NTGQDIMISV YNIHRSEVW
 451 EKAEEFLPER FDIDGAIPNE TNTDFKFIPF SGGPRKCVGD QFALMEAIVA
 501 LAVFLQRLNV ELVPDQTISM TTGATIHTTN GLYMKVSQR
 CHITS AT: 200-276, 347-352

RELATED SEQUENCES AVAILABLE WITH SEQLINK

MF Unspecified

CI MAN

SR CA

LC STN Files: CA, CAPLUS, USPATFULL

DT.CA Caplus document type: Patent

RL.P Roles from patents: BIOL (Biological study); PREP (Preparation); PRP (Properties)

1 REFERENCES IN FILE CA (1907 TO DATE)

1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

L4 ANSWER 2 OF 5 REGISTRY COPYRIGHT 2006 ACS on STN

RN 278251-53-9 REGISTRY

CN Protein (Arabidopsis thaliana clone ARATH-23APR03-C3001_1.p) (9CI) (CA INDEX NAME)

OTHER NAMES:

CN 2317: PN: US20040216190 SEQID: 6317 claimed protein

FS PROTEIN SEQUENCE

SQL 560

PATENT ANNOTATIONS (PNTE):

Sequence | Patent

Source | Reference

=====+=====

Not Given | US2004216190

| claimed

| SEQID 6317

SEQ 1 SPPPPRESQI FLHGSFEKRS SMESSLFSPS SSSYSSLFTA KPTRLSPKP
 51 KFTFSIRSSI EKPKPKLETN SSKSQSWVSP DWLTTLTRTL SSGKNDESGI
 101 PIANAKLDDV ADLLGGALFL PLYKWMNEYG PIYRLAAGPR NFVIVSDPAI
 151 AKHVLARNYPK YAKGLVAEVS EFLFGSGFAI AEGPLWTARR RAVVPSLHRR
 201 YLSVIVERVF CKCAERLVEK ~~QPYAEDGSA V NMEAKFSQMT LDVIGLSLFN~~
 251 ~~NYNFDSLTTD SPVIEAVYTA LKEAELRSTD LPYWKIDAL CKIVPROVKA~~
 301 EKAU TLIRET VEDLIAKCKE IVEREGERIN DEEYVNDADP SILRFLLASR
 351 EEVSSVQLRD DLLSMLVAGH ETTGSVLTWT LYLLSKNSSA LRKAQEEVDR
 401 VLEGRNPAFE DIKELKYITR CINESMRLYP HPPVLIRRAQ VPDILPGNYK
 451 VNTGQDIMIS VYNIHRSEV WEKAEEFLPE RFDIDGAIPN ETNTDFKFIP
 501 FSGGPRKCVG DQFALMEAIV ALAVFLQRLN VELVPDQTIS MTTGATIHTT
 551 NGLYMKVSQR

HITS AT: 221-297, 368-373

MF Unspecified

SQL 560

SEQ 201 YLSVIVERVF CKCAERLVEK LQPYAEDGSA VNMEAKFSQM
TLDVIGLSLF

=====

251 NYNFDLTTD SPVIEAVYTA LKEAELRSTD LPPYWKIDAL
CKIVPRQVKA

=====

351 EEVSSVQLRD DLLSMLVAGH ETTGSVLTWT LYLLSKNSSA
LRKAQEEVDR

====

HITS AT: 221-297, 368-373
LC STN Files: CA, CAPLUS

L6 ANSWER 10 OF 15 REGISTRY COPYRIGHT 2006 ACS on STN
RN 662363-73-7 REGISTRY
CN Protein (corn clone UC-ZMFLM017009E12_FLI.pep
fragment) (9CI) (CA
INDEX

NAME)
OTHER NAMES:
CN 1634: PN: US20040034888 SEQID: 70634 claimed protein
SQL 382

SEQ 151 DKVFCKCAER LIDKLEPYAL SGEPVNMEAR FSQLTLDVIG
LSLFFNYNFDS

=====

201 LTTDSPVIDA VYTALKEAEL RSTDLLPYWK VGFLCKIIPR
QIKAENAVTI

=====

301 QLRDDL SML VAGHETG SV LTWTIYLLSK DPTALRRAQD
EVDRVLQGR

=====

HITS AT: 191-216, 312-317
LC STN Files: CA, CAPLUS

L6 ANSWER 11 OF 15 REGISTRY COPYRIGHT 2006 ACS on STN
RN 636098-44-7 REGISTRY
CN GenBank AAR83120 (9CI) (CA INDEX NAME)

OTHER NAMES:
CN Chloroplast carotenoid epsilon-ring hydroxylase
(Arabidopsis
thaliana gene
CYP97C1)
CN GenBank AAR83120 (Translated from: GenBank AY424805)
SQL 539

SEQ 151 FLFGSGFAIA EGPLWTARRR AVVPSLHRRY LSVIVERVFC
KCAERLVEKL

201 QPYAEDGSAV NMEAKFSQMT LDVIGLSLFN YNFDSLTTDS
PVIEAVYTAL

=====

251 KEAELRSTD L PYWKIDALC KIVPRQVKAE KAVTLIRETV
EDLIAKCKEI

=====

301 VEREGERIND EEYVNDADPS ILRFLASRE EVSSVQLRDD
LLSMLVAGHE

351 TTGSVLTWTL YLLSKNSSAL RKAQEEVDRV LEGRNPAFED
IKELKYITRC

==

HITS AT: 200-276, 347-352

RELATED SEQUENCES AVAILABLE WITH SEQLINK

LC STN Files: CA, CAPLUS

L6 ANSWER 12 OF 15 REGISTRY COPYRIGHT 2006 ACS on STN
RN 536817-67-1 REGISTRY
CN Cytochrome P450 monooxygenase (Oryza sativa japonica
gene

OSJNBa0001014.16) (9CI) (CA INDEX NAME)

OTHER NAMES:

CN GenBank AAP54891

CN GenBank AAP54891 (Translated from: GenBank AE017117)

SQL 584

SEQ 201 AERLVEKLET SALSGKPVNM EARFSQMTLD VIGLSLFNYN
FDSLTS DSPV

251 IDAVYTALKE AELRSTDLLP YWKIDLLCKI VPRQIKA EKA
VNIIRNTVED

=====

351 SMLVAGHETT GSVLTWTIYL LSKDPAALRR AQAEVDRV LQ
GRLPRYEDLK

=====

HITS AT: 234-259, 355-360

RELATED SEQUENCES AVAILABLE WITH SEQLINK

LC STN Files: CA, CAPLUS

L6 ANSWER 13 OF 15 REGISTRY COPYRIGHT 2006 ACS on STN
RN 486504-98-7 REGISTRY

CI MAN
 SR CA
 LC STN Files: CA, CAPLUS
 DT.CA Caplus document type: Patent
 RL.P Roles from patents: BIOL (Biological study); PRP (Properties); USES
 (Uses)

1 REFERENCES IN FILE CA (1907 TO DATE)
 1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

L4 ANSWER 3 OF 5 REGISTRY COPYRIGHT 2006 ACS on STN
 RN 636098-44-79 REGISTRY
 CN GenBank AAR83120 (9CI) (CA INDEX NAME)
 OTHER NAMES:
 CN Chloroplast carotenoid epsilon-ring hydroxylase (Arabidopsis thaliana gene CYP97C1)
 CN GenBank AAR83120 (Translated from: GenBank AY424805)
 FS PROTEIN SEQUENCE
 SQL 539

SEQ 1 MESSLFSPSS SSYSSLFTAK PTRLLSPKPK FTFSIRSSIE KPKPKLETNS
 51 SKSQSWVSPD WLTTLTRTLS SGKNDESGIP IANAKLDDVA DLLGGALFLP
 101 LYKWMNEYGP IYRLAAGPRN FVIVSDPAIA KHVLRNYPKY AKGLVAEVSE
 151 FLFGSGFAIA EGPLWTARRR AVVPSLHRRY LSVIVERVFC KCAERLVEKLE
 201 ~~OPYAEDGSAV NMEAKESQMT LDVIGLSLFN YNFDSLTTDS PVIEAVYTAL~~
 =====
 251 ~~KEAELRSTDL LPYWKIDALC KIVPRQVKA E KAVTLIRETV EDLIAKCKEI~~
 =====
 301 VEREGERIND EEYVNDADPS ILRFLASRE EVSSVQLRDD LLSMLVAGHE
 351 TTGSVLTWTL YLLSKNSSAL RKAQEEVDRV LEGRNPAFED IKELKYITRC
 401 INESMRLYPH PPVLIRRAQV PDILPGNYKV NTGQDIMISV YNIHRSSEVW
 451 EKAEFLPER FDIDGAIPNE TNTDFKFIPF SGGPRKCVGD QFALMEAIWA
 501 LAVFLQRLNV ELVPDQTISM TTGATIHTTN GLYMKVSQR

HITS-AT: 200-276, 347-352

RELATED SEQUENCES AVAILABLE WITH SEQLINK

MF Unspecified
 CI MAN
 SR GenBank
 LC STN Files: CA, CAPLUS
 DT.CA Caplus document type: Journal
 RL.NP Roles from non-patents: BIOL (Biological study); PRP (Properties)
 1 REFERENCES IN FILE CA (1907 TO DATE)
 1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

L4 ANSWER 4 OF 5 REGISTRY COPYRIGHT 2006 ACS on STN
 RN 486504-98-79 REGISTRY
 CN GenBank CAB64216 (9CI) (CA INDEX NAME)
 OTHER NAMES:
 CN GenBank CAB64216 (Translated from: GenBank AL132958)
 FS PROTEIN SEQUENCE
 SQL 566

SEQ 1 MESSLFSPSS SSYSSLFTAK PTRLLSPKPK FTFSIRSSIE KPKPKLETNS
 51 SKSQSWVSPD WLTTLTRTLS SGKNDESGIP IANAKLDDVA DLLGGALFLP
 101 LYKWMNEYGP IYRLAAGPRN FVIVSDPAIA KHVLRNYPKY AKGLVAEVSE
 151 FLFGSGFAIA EGPLWTVISS PPISILKFLE LWKRRVAVPS LHRRYLSVIV
 201 ERVFCCKAER LVEKLQPYAE DGSANMEAK FSQMTLDVIG LSLFNYNFDS*
 =====
 251 ~~TTDSPVIEA VYTALKEAEL RSTDLLPYWK ASFLCFFCGL LIIDALCKIV~~

*Gen Bank record
 printed at end
 of search*

CN GenBank CAB64216 (9CI) (CA INDEX NAME)
OTHER NAMES:
CN GenBank CAB64216 (Translated from: GenBank AL132958)
SQL 566

SEQ 201 ERVFCKCAER LVEKLQPYAE DGSAVNMEAK FSQMTLDVIG
LSLNFYNFDS

=====

251 LTTDSPVIEA VYTALKEAEL RSTDLLPYWK ASFLCFFCGL
LIIDALCKIV

=====

351 FLLASREEVS SVQLRDDLLS MLVAGHETTG SVLTWTLYLL
SKNSSLARKA

=====

HITS AT: 215-266, 374-379

L6 ANSWER 14 OF 15 REGISTRY COPYRIGHT 2006 ACS on STN
RN 482144-13-8 REGISTRY
CN GenBank AAK20054 (9CI) (CA INDEX NAME)
OTHER NAMES:
CN 72: PN: US20050150002 PAGE: 4 claimed protein
CN GenBank AAK20054 (Translated from: GenBank AC025783)
SQL 584

SEQ 201 AERLVEKLET SALSGKPVNM EARFSQMTLD VIGLSLNFYN
FDSLTSDSPV

=====

251 IDAVYTALKE AELRSTDLLP YWKIDLLCKI VPRQIKAeka
VNIIRNTVED

=====

351 SMLVAGHETT GSVLTWTIYL LSKDPAALRR AQAQVDRVLQ
GRLPRYEDLK

=====

HITS AT: 234-259, 355-360

RELATED SEQUENCES AVAILABLE WITH SEQLINK

LC STN Files: CA, CAPLUS, USPATFULL

L6 ANSWER 15 OF 15 REGISTRY COPYRIGHT 2006 ACS on STN
RN 437167-31-2 REGISTRY
CN Cytochrome P450 (Arabidopsis thaliana clone RAFL09-78-
C14 (R19513)
gene

At3g53130) (9CI) (CA INDEX NAME)
OTHER NAMES:
CN GenBank AAM13903
CN GenBank AAM13903 (Translated from: GenBank AY091083)

```

=====
301 PRQVKA EKAV TLIRETVEDL IAKCKEIVER EGERINDEEY VNDADPSILR
351 FLLASREEVS SVQLRDDLLS MLVAGHETTG SVLTWTLYLL SKNSSALRKA
401 QEEVDRLVLE RNP AFEDIKE LKYITRCINE SMRLYPHPV LIRRAQVPDI
451 LPGNYKVNTG QDIMISVYNI HRSSEVWEKA EEFLPERFDI DGAIPNETNT
501 DFKFIPFSGG PRKCVGDQFA LMEAIVALAV FLQRLNVELV PDQTISMTTG
551.ATIHTTNGLY MKVSQR

```

HITS AT: 215-268 , 374-379

MF Unspecified

CI MAN

SR GenBank

L4 ANSWER 5 OF 5 REGISTRY COPYRIGHT 2006 ACS on STN

RN ~~437167-31-2s~~ REGISTRY

CN Cytochrome P450 (Arabidopsis thaliana clone RAFL09-78-C14 (R19513) gene At3g53130) (9CI) (CA INDEX NAME)

OTHER NAMES:

CN GenBank AAM13903

CN GenBank AAM13903 (Translated from: GenBank AY091083)

FS PROTEIN SEQUENCE

SQL 552

```

SEQ      1 QIFLHGSFEK RSSMESSLFS PSSSSYSSLF TAKPTRLLSP KPKFTFSIRS
        51 SIEKPKPKLE TNSSKSQSWV SPDWLTTLTR TLSSGKNDES GIPIANAKLD
       101 DVADLLGGAL FLPLYKWMNE YGPIYRLAAG PRNFVIVSDP AIAKHVLRNY
       151 PKYAKGLVAE VSEFLFGSGF AIAEGPLWTA RRRAVVPSLH RRYLSVIVER
       201 VFCKCAERLV EKLPYAEDG SAVNMEAKFS QMTLDVIGLS LFNYNFDSLT
              =====
       251 TDSPVIEAVY TALKEAELRS TDLLPYWKID ALCKIVPROV KAEKAVTLIR
              =====
       301 ETVEDLIAKC KEIVEREGE INDEEYVND A DPSILRFLA SREEVSSVQL
       351 RDDLLSMLVA GHETTGSVLT WTLYLLSKNS SALRKAQEEV DRVLEGRNPA
       401 FEDIKELKYI TRCINESMRL YPHPPVLIRR AQVPDILPGN YKVNTGQDIM
       451 ISVYNIHRSS EVWEKAEFL PERFDIDGAI PNETNTDFKF IPFSGGPRKC
       501 VGDQFALMEA IVALAVFLQR LNVELVPDQT ISMTTGATIH TTNGLYMKVS
       551 QR

```

HITS AT: 213-289 , 360-365

MF Unspecified

CI MAN

SR CA

LC STN Files: CA, CAPLUS

DT.CA Caplus document type: Journal

RL.NP Roles from non-patents: BIOL (Biological study); PRP (Properties)

1 REFERENCES IN FILE CA (1907 TO DATE)

1 REFERENCES IN FILE CAPLUS (1907 TO DATE)

=> fil capl uspatf; s l4

~~FILE 'CAPLUS'~~ ENTERED AT 15:12:41 ON 09 MAR 2006

USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT.

PLEASE SEE "HELP USAGETERMS" FOR DETAILS.

COPYRIGHT (C) 2006 AMERICAN CHEMICAL SOCIETY (ACS)

~~FILE 'USPATFULL'~~ ENTERED AT 15:12:41 ON 09 MAR 2006

CA INDEXING COPYRIGHT (C) 2006 AMERICAN CHEMICAL SOCIETY (ACS)

~~d5~~-----5-L4

*Reg. 3 try answer
set crossed into
bibliographic
files to get
references*

=> ~~dup rem l5~~

PROCESSING COMPLETED FOR L5

~~L6~~ ~~4-DUP REM L5 (1 DUPLICATE REMOVED)~~
ANSWERS '1-4' FROM FILE CAPLUS

=> ~~d ibib ed abs hitrn 1-4~~; fil hom

L6 ANSWER 1 OF 4 CAPLUS COPYRIGHT 2006 ACS on STN DUPLICATE 1
 ACCESSION NUMBER: 2005:592143 CAPLUS
 DOCUMENT NUMBER: 143:110628
 TITLE: Cloning and sequences of plant cytochrome P
 450-dependent carotenoid hydroxylases for use in
 engineering carotenoid metabolism in plants
 INVENTOR(S): Dellapenna, Dean; Tian, Li; Kim, Joonyul
 PATENT ASSIGNEE(S): USA
 SOURCE: U.S. Pat. Appl. Publ., 135 pp.
 CODEN: USXXCO
 DOCUMENT TYPE: Patent
 LANGUAGE: English
 FAMILY ACC. NUM. COUNT: 1
 PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 2005150002	A1	20050707	US 2004-751235	20040102
WO 2005067512	A2	20050728	WO 2004-US44033	20041229

W: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH,
 CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD,
 GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC,
 LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI,
 NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY,
 TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW
 RW: BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW, AM,
 AZ, BY, KG, KZ, MD, RU, TJ, TM, AT, BE, BG, CH, CY, CZ, DE, DK,
 EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, MC, NL, PL, PT,
 RO, SE, SI, SK, TR, BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML,
 MR, NE, SN, TD, TG

PRIORITY APPLN. INFO.: US 2004-751235 A 20040102

ED Entered STN: 08 Jul 2005

AB The present invention relates to genes, proteins and methods comprising
 carotenoid monooxygenases in the cytochrome P 450 family. In a preferred
 embodiment, the present invention relates to altering carotenoid ratios in
 plants and microorganisms using LUT1 ϵ -hydroxylases and/or CYP97A
 β -hydroxylases. The nucleotide sequences and the encoded amino acid
 sequences of LUT1 ϵ -hydroxylases and CYP97A β -hydroxylases
 from various plants are provided.

IT ~~857687=96-8P~~ Use Registry # to match reference to sequence

RL: BPN (Biosynthetic preparation); BSU (Biological study, unclassified);
 PRP (Properties); BIOL (Biological study); PREP (Preparation)
 (amino acid sequence; cloning and sequences of plant cytochrome P
 450-dependent carotenoid hydroxylases for use in engineering carotenoid
 metabolism in plants)

L6 ANSWER 2 OF 4 CAPLUS COPYRIGHT 2006 ACS on STN
 ACCESSION NUMBER: 2004:930169 CAPLUS
 DOCUMENT NUMBER: 141:361551
 TITLE: Nucleic acid molecules and encoded proteins associated
 with plants and their uses for plant improvement
 INVENTOR(S): Kovalic, David K.
 PATENT ASSIGNEE(S): USA

SOURCE: U.S. Pat. Appl. Publ., 14 pp., Cont.-in-part of U.S. Ser. No. 424,599.
 CODEN: USXXCO
 DOCUMENT TYPE: Patent
 LANGUAGE: English
 FAMILY ACC. NUM. COUNT: 76
 PATENT INFORMATION:

Kovalitz →

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 2004216190	A1	20041028	US 2003-739930	20031218
US 2004031072	A1	20040212	US 2003-424599	20030428
US 2004214272	A1	20041028	US 2003-425115	20030428
US 2004216190	A1	20041028	US 2003-739930	20031218
PRIORITY APPLN. INFO.:			US 2003-424599	A2 20030428
			US 2003-425115	A2 20030428
			US 2003-739930	A 20031218
			US 1999-304517	B1 19990506
			US 2001-985678	B2 20011105

ED Entered STN: 06 Nov 2004

AB Recombinant polynucleotides useful for improvement of plants are provided. In particular, a total of 5544 cDNA sequences are provided from cDNA libraries generated from Arabidopsis thaliana, Brassica napus (rape), Zea mays (corn), Glycine max (soybean), and Triticum aestivum (wheat). The polypeptides encoded by these polynucleotide sequences are also provided. The open reading frame in each polynucleotide sequence is identified by a combination of predictive and homol. based methods. Functions of polypeptides are determined using a hierarchical classification tool (FunCAT) and five public classification schemes (GO BP, GO CC, GO MF, KEGG, and EC) and one internal Monsanto classification scheme (POI). The disclosed recombinant polynucleotides and polypeptides find use in production of transgenic plants to produce plants having improved properties. [This abstract record is one of three records for this document necessitated by the large number of index entries required to fully index the document and publication system constraints.].

IT ~~778251=53=9~~

RL: BSU (Biological study, unclassified); BUU (Biological use, unclassified); PRP (Properties); BIOL (Biological study); USES (Uses) (amino acid sequence; nucleic acid mols. and encoded proteins associated with plants and their uses for plant improvement)

L6 ANSWER 3 OF 4 CAPLUS COPYRIGHT 2006 ACS on STN

ACCESSION NUMBER: 2004:39274 CAPLUS

DOCUMENT NUMBER: 140:194233

TITLE: The Arabidopsis LUT1 locus encodes a member of the cytochrome P450 family that is required for carotenoid ϵ -ring hydroxylation activity

AUTHOR(S): Tian, Li; Musetti, Valeria; Kim, Joonyul; Magallanes-Lundback, Maria; DellaPenna, Dean

CORPORATE SOURCE: Department of Biochemistry and Molecular Biology, Michigan State University, East Lansing, MI, 48824, USA

SOURCE: Proceedings of the National Academy of Sciences of the United States of America (2004), 101(1), 402-407
 CODEN: PNASA6; ISSN: 0027-8424

PUBLISHER: National Academy of Sciences

DOCUMENT TYPE: Journal

LANGUAGE: English

ED Entered STN: 16 Jan 2004

AB Lutein, a dihydroxy xanthophyll, is the most abundant carotenoid in plant

photosynthetic tissues and plays crucial structural and functional roles in the light-harvesting complexes. Carotenoid β - and ϵ -hydroxylases catalyze the formation of lutein from α -carotene (β , ϵ -carotene). In contrast to the well studied β -hydroxylases that have been cloned and characterized from many organisms, the ϵ -hydroxylase has only been genetically defined by the *lut1* mutation in Arabidopsis. We have isolated the *LUT1* gene by positional cloning and found that, in contrast to all known carotenoid hydroxylases, which are the nonheme diiron monooxygenases, *LUT1* encodes a cytochrome P 450-type monooxygenase, CYP97C1. Introduction of a null mutant allele of *LUT1*, *lut1-3*, into the β -hydroxylase 1/ β -hydroxylase 2 (*b1 b2*) double-mutant background, in which both Arabidopsis β -hydroxylases are disrupted, yielded a genotype (*lut1-3 b1 b2*) in which all three known carotenoid hydroxylase activities are eliminated. Surprisingly, hydroxylated β -rings were still produced in *lut1-3 b1 b2*, suggesting that a fourth unknown carotenoid β -hydroxylase exists in vivo that is structurally unrelated to β -hydroxylase 1 or 2. A second chloroplast-targeted member of the CYP97 family, CYP97A3, is 49% identical to *LUT1* and hypothesized as a likely candidate for this addnl. β -ring hydroxylation activity. Overall, *LUT1* defines a class of carotenoid hydroxylases that has evolved independently from and uses a different mechanism than nonheme diiron β -hydroxylases.

IT ~~636098-44-7~~ GenBank AAR83120

RL: BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)

(amino acid sequence; Arabidopsis *LUT1* locus encodes member of the cytochrome P 450 family that is required for carotenoid ϵ -ring hydroxylation activity)

REFERENCE COUNT: 30 THERE ARE 30 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

L6 ANSWER 4 OF 4 CAPLUS COPYRIGHT 2006 ACS on STN

ACCESSION NUMBER: 2002:280980 CAPLUS

DOCUMENT NUMBER: 137:28849

TITLE: Functional annotation of a full-length Arabidopsis cDNA collection

AUTHOR(S): Seki, Motoaki; Narusaka, Mari; Kamiya, Asako; Ishida, Junko; Satou, Masakazu; Sakurai, Tetsuya; Nakajima, Maiko; Enju, Akiko; Akiyama, Kenji; Oono, Youko; Muramatsu, Masami; Hayashizaki, Yoshihide; Kawai, Jun; Carninci, Piero; Itoh, Masayoshi; Ishii, Yoshiyuki; Arakawa, Takahiro; Shibata, Kazuhiro; Shinagawa, Akira; Shinozaki, Kazuo

CORPORATE SOURCE: Plant Mutation Exploration Team, Plant Functional Genomics Res. Group, RIKEN Genomic Sciences Center (GSC), 3-1-1 Koyadai, Tsukuba, 305-0074, Japan

SOURCE: Science (Washington, DC, United States) (2002), 296(5565), 141-145

CODEN: SCIEAS; ISSN: 0036-8075

PUBLISHER: American Association for the Advancement of Science

DOCUMENT TYPE: Journal

LANGUAGE: English

ED Entered STN: 16 Apr 2002

AB Full-length cDNAs are essential for the correct annotation of genomic sequences and for the functional anal. of genes and their products. About 155,144 RIKEN Arabidopsis full-length (RAFL) cDNA clones were isolated. The 3'-end expressed sequence tags (ESTs) of 155,144 RAFL cDNAs were clustered into 14,668 nonredundant cDNA groups, about 60% of predicted genes. 5'-ESTs were also obtained from 14,034 nonredundant cDNA groups

and a promoter database constructed. The sequence database of the RAFL cDNAs is useful for promoter anal. and correct annotation of predicted transcription units and gene products. Furthermore, the full-length cDNAs are useful resources for analyses of the expression profiles, functions, and structures of plant proteins. [This abstract record is one of sixteen records for this document necessitated by the large number of index entries required to fully index the document and publication system constraints.].



IT ~~437167=31-2~~

RL: BSU (Biological study, unclassified); PRP (Properties); BIOL
(Biological study)
(amino acid sequence; functional annotation of a full-length
Arabidopsis cDNA collection)

REFERENCE COUNT: 27 THERE ARE 27 CITED REFERENCES AVAILABLE FOR THIS
RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

FILE 'HOME' ENTERED AT 15:12:55 ON 09 MAR 2006

=>



My NCBI [\[Sign In\]](#) [\[Register\]](#)

[PubMed](#) [Nucleotide](#) [Protein](#) [Genome](#) [Structure](#) [PMC](#) [Taxonomy](#) [OMIM](#) [Books](#)

Search for

Limits Preview/Index History Clipboard Details

Display Show Send to

Range: from to Features: ☐ SNP ☒ CDD ☒ MGC ☒ HPRD ☒ STS ☒ tRNA

☐ 1: [CAB64216](#). Reports Cytochrom P450-li...[gi:6630733]

BLink, Conserved
Domains, Links

Comment Features Sequence

LOCUS CAB64216 566 aa linear PLN 16-APR-2005
DEFINITION Cytochrom P450-like protein [Arabidopsis thaliana].
ACCESSION CAB64216
VERSION CAB64216.1 GI:6630733
DBSOURCE embl locus ATT4D2, accession [AL132958.1](#)
KEYWORDS .
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM [Arabidopsis thaliana](#)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1
AUTHORS Nyakatura,G., Fartmann,B., Dauner,D., Sterr,W., Holland,R.,
Weichselgartner,M., Mewes,H.W., Lemcke,K., Mayer,K.F.X., Quetier,F.
and Salanoubat,M.
JOURNAL Unpublished
REFERENCE 2 (residues 1 to 566)
AUTHORS EU Arabidopsis sequencing,project.
TITLE Direct Submission
JOURNAL Submitted (21-DEC-1999) MIPS, at the Max-Planck-Institut fuer
Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:
lemcke@mips.biochem.mpg.de,mayer@mips.biochem.mpg.de Project
Coordinator: Marcel Salanoubat and Francis Quetier, Groupement
d'Interet Public, Centre National de Sequencage - GENOSCOPE; 2 rue
Gaston Cremieux, BP191, 91006 Evry Cedex, France;
http://www.genoscope.cns.fr
COMMENT Information on performance of analysis and a more detailed
annotation of this entry and other sequences of chromosomes 3, 4
and 5 can be viewed at: <http://www.mips.biochem.mpg.de/proj/thal/>.
FEATURES Location/Qualifiers
source 1..566
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/chromosome="3"
/ecotype="Columbia"
Protein 1..566
/product="Cytochrom P450-like protein"
Region 101..545
/region_name="Cytochrome P450"
/note="p450"
/db_xref="CDD:25378"
CDS 1..566

```
/gene="T4D2.60"  
/coded_by="join(AL132958.1:16880..17427,  
AL132958.1:17532..17859,AL132958.1:17913..18110,  
AL132958.1:18200..18301,AL132958.1:18400..18540,  
AL132958.1:18610..18717,AL132958.1:18879..18961,  
AL132958.1:19059..19221,AL132958.1:19317..19346)"  
/note="strong similarity to CYTOCHROME P450 97B2, Glycine  
max., gb:O48921  
Contains Cytochrome P450 cysteine heme-iron ligand  
signature AA507-516"  
/db_xref="GOA:Q9SCP8"  
/db_xref="InterPro:IPR001128"  
/db_xref="InterPro:IPR002401"  
/db_xref="UniProtKB/TrEMBL:Q9SCP8"
```

ORIGIN

```
1 messlfpss ssysslftak ptrllspkpk ftfsirssie kpkpkletns sksqswvspd  
61 wlttltrtlls sgkndesgip ianaklddva dllggalflp lykwmneygp iyrlaagprn  
121 fvivsdpaia khvlrnypky akglvaevse flfgsgfaia egplwtviss ppisilkfle  
181 lwkrravvps lhrrylsviv ervfckcaer lveklfpayae dgsavameak fsgmtldvlg  
241 lskkmynfde ltttdspviea vytalkaeael rstddlpywk asflcfcgcl liidalckiv  
301 prqvkaekav tliretvedl iakckeiver egerindeey vndadpsilr fllasreevs  
361 svqlrddlls mlvaghettg svltwtlyll sknssalrka qeevdrvleg rnpafedike  
421 lkyitrcine smrlyphppv lirraqvpdi lpgnykvntg qdimisvyni hrssevweka  
481 eeelfperfdi dgaipnetnt dfkfipfsgg prkcvgdqfa lmeaivalav flqrlnvelv  
541 pdqtismttg atihttngly mkvsqr
```

//

[Disclaimer](#) | [Write to the Help Desk](#)
[NCBI](#) | [NLM](#) | [NIH](#)

Feb 1 2006 13:21:03

=> d his nofile

(FILE 'HOME' ENTERED AT 15:06:46 ON 09 MAR 2006)
D SAVED

FILE 'REGISTRY' ENTERED AT 15:07:25 ON 09 MAR 2006
ACT WOR235SEQ/A

L1 14471 SEA ABB=ON [AG]G.[DE]T[TS]/SQSP

FILE 'LREGISTRY' ENTERED AT 15:07:34 ON 09 MAR 2006

L2 0 SEA ABB=ON LQPYAEDGSAVNMEAKFSQMTLDVIGLSLFN|VYTALKEAELRSTDLLPYW
KIDALCKIVPRQ|VIGLSLFNYNFDSLTTDSPVIEAVYTALKEA/SQSP

FILE 'REGISTRY' ENTERED AT 15:11:24 ON 09 MAR 2006

L3 6 SEA ABB=ON LQPYAEDGSAVNMEAKFSQMTLDVIGLSLFN|VYTALKEAELRSTDLLPYW
KIDALCKIVPRQ|VIGLSLFNYNFDSLTTDSPVIEAVYTALKEA/SQSP
SAVE TEMP L3 WOR235SEQ2/A

L4 5 SEA ABB=ON L1 AND L3

FILE 'REGISTRY' ENTERED AT 15:12:09 ON 09 MAR 2006

D QUE L4
D RN CN SQL KWIC NTE LC 1-5

FILE 'CAPLUS, USPATFULL' ENTERED AT 15:12:41 ON 09 MAR 2006

L5 5 SEA ABB=ON L4
L6 4 DUP REM L5 (1 DUPLICATE REMOVED)
ANSWERS '1-4' FROM FILE CAPLUS
D IBIB ED ABS HITRN 1-4

FILE 'HOME' ENTERED AT 15:12:55 ON 09 MAR 2006
D SAVED

FILE 'REGISTRY' ENTERED AT 15:13:18 ON 09 MAR 2006
SAVE TEMP L4 WOR235SEQ3/A

=>

SUBSET OF STIC SEQUENCE SEARCH

Worley 10/751235

for SEQ ID NO:

Page 5

12

=> dup rem L5

PROCESSING COMPLETED FOR L5

L6 4 DUP REM L5 (1 DUPLICATE REMOVED)
ANSWERS '1-4' FROM FILE CAPLUS

=> d. bib ed abs hitrn 1-4; fil hom

L6 ANSWER 1 OF 4 CAPLUS COPYRIGHT 2006 ACS on STN DUPLICATE 1

ACCESSION NUMBER: 2005:592143 CAPLUS

DOCUMENT NUMBER: 143:110628

TITLE: Cloning and sequences of plant cytochrome P
450-dependent carotenoid hydroxylases for use in
engineering carotenoid metabolism in plants
Dellapenna, Dean; Tian, Li; Kim, Joonyul

INVENTOR(S):

PATENT ASSIGNEE(S): USA

SOURCE: U.S. Pat. Appl. Publ., 135 pp.

CODEN: USXXCO

DOCUMENT TYPE: Patent

LANGUAGE: English

FAMILY ACC. NUM. COUNT: 1

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 2005150002	A1	20050707	US 2004-751235	20040102
WO 2005067512	A2	20050728	WO 2004-US44033	20041229
W:	AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW			
RW:	BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR, BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG			

PRIORITY APPLN. INFO.:

US 2004-751235 A 20040102

ED Entered STN: 08 Jul 2005

AB The present invention relates to genes, proteins and methods comprising carotenoid monooxygenases in the cytochrome P 450 family. In a preferred embodiment, the present invention relates to altering carotenoid ratios in plants and microorganisms using LUT1 α -hydroxylases and/or CYP97A β -hydroxylases. The nucleotide sequences and the encoded amino acid sequences of LUT1 α -hydroxylases and CYP97A β -hydroxylases from various plants are provided.

IT 857687-96-8P Use Registry # to match reference to sequence

RL: BPN (Biosynthetic preparation); BSU (Biological study, unclassified);

PRP (Properties); BIOL (Biological study); PREP (Preparation)

(amino acid sequence; cloning and sequences of plant cytochrome P
450-dependent carotenoid hydroxylases for use in engineering carotenoid
metabolism in plants)

L6 ANSWER 2 OF 4 CAPLUS COPYRIGHT 2006 ACS on STN

ACCESSION NUMBER: 2004:930169 CAPLUS

DOCUMENT NUMBER: 141:361551

TITLE: Nucleic acid molecules and encoded proteins associated
with plants and their uses for plant improvement

INVENTOR(S): Kovalic, David K.

PATENT ASSIGNEE(S): USA

Searched by Barb O'Bryen, STIC 2-2518

SOURCE: U.S. Pat. Appl. Publ., 14 pp., Cont.-in-part of U.S. Ser. No. 424,599.
 CODEN: USXXCO
 DOCUMENT TYPE: Patent
 LANGUAGE: English
 FAMILY ACC. NUM. COUNT: 76
 PATENT INFORMATION:

Kowli →

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 2004216190	A1	20041028	US 2003-739930	20031218
US 2004031072	A1	20040212	US 2003-424599	20030428
US 2004214272	A1	20041028	US 2003-425115	20030428
US 2004216190	A1	20041028	US 2003-739930	20031218
PRIORITY APPLN. INFO.:			US 2003-424599	A2 20030428
			US 2003-425115	A2 20030428
			US 2003-739930	A 20031218
			US 1999-304517	B1 19990506
			US 2001-985678	B2 20011105

ED Entered STN: 06 Nov 2004

AB Recombinant polynucleotides useful for improvement of plants are provided. In particular, a total of 5544 cDNA sequences are provided from cDNA libraries generated from *Arabidopsis thaliana*, *Brassica napus* (rape), *Zea mays* (corn), *Glycine max* (soybean), and *Triticum aestivum* (wheat). The polypeptides encoded by these polynucleotide sequences are also provided. The open reading frame in each polynucleotide sequence is identified by a combination of predictive and homol. based methods. Functions of polypeptides are determined using a hierarchical classification tool (FunCAT) and five public classification schemes (GO_BP, GO_CC, GO_MF, KEGG, and EC) and one internal Monsanto classification scheme (POI). The disclosed recombinant polynucleotides and polypeptides find use in production of transgenic plants to produce plants having improved properties. [This abstract record is one of three records for this document necessitated by the large number of index entries required to fully index the document and publication system constraints.].

IT 778251-53-9 *

RL: BSU (Biological study, unclassified); BUU (Biological use, unclassified); PRP (Properties); BIOL (Biological study); USES (Uses) (amino acid sequence; nucleic acid mols. and encoded proteins associated with plants and their uses for plant improvement)

L6 ANSWER 3 OF 4 CAPLUS COPYRIGHT 2006 ACS on STN

ACCESSION NUMBER: 2004:39274 CAPLUS

DOCUMENT NUMBER: 140:194233

TITLE: The *Arabidopsis* LUT1 locus encodes a member of the cytochrome P450 family that is required for carotenoid ϵ -ring hydroxylation activity

AUTHOR(S): Tian, Li; Musetti, Valeria; Kim, Joonyul; Magallanes-Lundback, Maria; DellaPenna, Dean

CORPORATE SOURCE: Department of Biochemistry and Molecular Biology, Michigan State University, East Lansing, MI, 48824, USA

SOURCE: Proceedings of the National Academy of Sciences of the United States of America (2004), 101(1), 402-407
 CODEN: PNASA6; ISSN: 0027-8424

PUBLISHER: National Academy of Sciences

DOCUMENT TYPE: Journal

LANGUAGE: English

ED Entered STN: 16 Jan 2004

AB Lutein, a dihydroxy xanthophyll, is the most abundant carotenoid in plant

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 8, 2006, 11:24:53 ; Search time 165 Seconds
(without alignments)
25.323 Million cell updates/sec

Title: US-10-751-235-14
Perfect score: 33
Sequence: 1 FXGXGXXCXG 10

Scoring table: BLOSUM62DX *
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database: Published Applications AA Main:

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	100.0	10	3	US-09-739-254-129
2	33	100.0	10	3	US-09-904-615-129
3	33	100.0	10	4	US-10-067-534-15
4	33	100.0	10	4	US-10-210-965-5
5	33	100.0	10	4	US-10-054-988-129
6	33	100.0	10	4	US-10-236-433-11
7	33	100.0	10	5	US-10-477-526-7
8	33	100.0	10	5	US-10-477-526-8
9	33	100.0	10	5	US-10-751-235-14
10	33	100.0	10	5	US-10-751-235-15
11	33	100.0	11	4	US-10-097-559-28
12	33	100.0	11	5	US-10-804-772-26
13	33	100.0	11	5	US-10-804-772-29
14	33	100.0	14	4	US-10-373-877-10
15	33	100.0	14	4	US-10-373-877-19
16	33	100.0	14	4	US-10-373-877-20
17	33	100.0	14	4	US-10-373-877-21
18	33	100.0	14	4	US-10-373-877-22
19	33	100.0	14	4	US-10-373-877-23
20	33	100.0	14	4	US-10-373-877-24
21	33	100.0	14	4	US-10-373-877-25
22	33	100.0	14	5	US-10-865-478-774
23	33	100.0	15	3	US-09-957-674-13
24	33	100.0	15	4	US-10-378-1
25	33	100.0	15	5	US-10-882-241-17
26	33	100.0	16	4	US-10-163-198-90
27	33	100.0	16	5	US-10-926-683-1257

28	33	100.0	18	2	US-08-647-444-12	Sequence 12, Appl
29	33	100.0	20	5	US-10-855-595-6	Sequence 6, Appl
30	33	100.0	20	5	US-10-855-595-7	Sequence 7, Appl
31	33	100.0	20	5	US-10-855-595-8	Sequence 8, Appl
32	33	100.0	20	5	US-10-855-595-9	Sequence 9, Appl
33	33	100.0	20	5	US-10-855-595-10	Sequence 10, Appl
34	33	100.0	20	5	US-10-855-532-6	Sequence 6, Appl
35	33	100.0	20	5	US-10-855-532-7	Sequence 7, Appl
36	33	100.0	20	5	US-10-855-532-8	Sequence 8, Appl
37	33	100.0	20	5	US-10-855-532-9	Sequence 9, Appl
38	33	100.0	20	5	US-10-855-532-10	Sequence 10, Appl
39	33	100.0	21	4	US-10-021-425-25	Sequence 25, Appl
40	33	100.0	21	5	US-10-900-856-28	Sequence 28, Appl
41	33	100.0	23	4	US-10-103-196-31	Sequence 31, Appl
42	33	100.0	30	3	US-09-749-637A-126	Sequence 126, App
43	33	100.0	30	3	US-09-749-637A-129	Sequence 129, App
44	33	100.0	30	3	US-09-749-637A-283	Sequence 283, App
45	33	100.0	30	4	US-10-433-485A-1	Sequence 1, Appl
46	33	100.0	30	5	US-10-839-227-126	Sequence 126, App
47	33	100.0	30	5	US-10-839-227-129	Sequence 129, App
48	33	100.0	30	5	US-10-839-227-283	Sequence 283, App
49	33	100.0	32	4	US-10-437-963-159510	Sequence 159510,
50	33	100.0	35	4	US-10-351-641-487	Sequence 487, App
51	33	100.0	35	4	US-10-351-641-488	Sequence 488, App
52	33	100.0	35	4	US-10-351-641-489	Sequence 489, App
53	33	100.0	35	4	US-10-351-641-490	Sequence 490, App
54	33	100.0	35	4	US-10-351-641-491	Sequence 491, App
55	33	100.0	35	4	US-10-351-641-492	Sequence 492, App
56	33	100.0	35	4	US-10-351-641-493	Sequence 493, App
57	33	100.0	35	4	US-10-351-641-523	Sequence 523, App
58	33	100.0	35	4	US-10-351-641-524	Sequence 524, App
59	33	100.0	35	4	US-10-351-641-525	Sequence 525, App
60	33	100.0	35	4	US-10-351-641-526	Sequence 526, App
61	33	100.0	35	4	US-10-351-641-527	Sequence 527, App
62	33	100.0	35	4	US-10-351-641-528	Sequence 528, App
63	33	100.0	35	4	US-10-351-641-529	Sequence 529, App
64	33	100.0	39	4	US-10-147-140-54	Sequence 54, Appl
65	33	100.0	40	3	US-09-894-882-348	Sequence 348, App
66	33	100.0	40	3	US-09-894-882-482	Sequence 482, App
67	33	100.0	40	5	US-10-894-314A-348	Sequence 348, App
68	33	100.0	40	5	US-10-894-314A-482	Sequence 482, App
69	33	100.0	41	4	US-10-425-115-24467	Sequence 24467,
70	33	100.0	42	4	US-10-424-599-143366	Sequence 143366,
71	33	100.0	42	4	US-10-424-599-256014	Sequence 256014,
72	33	100.0	42	4	US-10-425-115-238482	Sequence 238482,
73	33	100.0	43	4	US-10-424-599-187772	Sequence 187772,
74	33	100.0	44	4	US-10-424-599-271553	Sequence 271553,
75	33	100.0	44	4	US-10-425-115-239272	Sequence 239272,
76	33	100.0	46	4	US-10-267-682-208	Sequence 208, App
77	33	100.0	46	4	US-10-267-748-208	Sequence 208, App
78	33	100.0	49	4	US-10-424-599-157117	Sequence 157117,
79	33	100.0	49	4	US-10-425-115-238518	Sequence 238518,
80	33	100.0	49	4	US-10-425-115-240869	Sequence 240869,
81	33	100.0	49	4	US-10-425-115-324478	Sequence 324478,
82	33	100.0	49	4	US-10-425-115-351931	Sequence 351931,
83	33	100.0	50	4	US-10-425-115-260285	Sequence 260285,
84	33	100.0	51	4	US-10-424-599-167757	Sequence 167757,
85	33	100.0	51	4	US-10-424-599-177489	Sequence 177489,
86	33	100.0	51	4	US-10-424-599-214094	Sequence 214094,
87	33	100.0	52	3	US-09-864-761-45740	Sequence 45740, A
88	33	100.0	52	4	US-10-424-599-254214	Sequence 254214,
89	33	100.0	52	5	US-10-946-683-1061	Sequence 1061, Ap
90	33	100.0	54	4	US-10-424-599-273991	Sequence 273991,
91	33	100.0	54	4	US-10-437-963-185080	Sequence 185080,
92	33	100.0	55	3	US-09-864-761-34668	Sequence 34668, A
93	33	100.0	55	4	US-10-106-698-6691	Sequence 6691, Ap
94	33	100.0	55	4	US-10-425-115-250988	Sequence 250988,
95	33	100.0	56	4	US-10-425-115-361393	Sequence 361393,
96	33	100.0	57	4	US-10-424-599-199843	Sequence 199843,
97	33	100.0	57	4	US-10-425-115-261158	Sequence 261158,
98	33	100.0	59	4	US-10-424-599-176580	Sequence 176580,
99	33	100.0	59	4	US-10-425-115-280265	Sequence 280265,
100	33	100.0	59	4	US-10-425-115-314843	Sequence 314843,

ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_82046C.1.p
US-10-424-599-265810

Query Match 100.0%; Score 122; DB 4; Length 222;
Best Local Similarity 100.0%; Pred. No. 3.7e-11;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LVAEVSEFLFGSGFAIAGPLWTA 24
|||
DB 157 LVAEVSEFLFGSGFAIAGPLWTA 180
|||

RESULT 3

US-10-751-235-4
Sequence 4, Application US/10751235
Publication No. US20050150002A1
GENERAL INFORMATION:

APPLICANT: Dellapenna, Dean
APPLICANT: Tian, Li
APPLICANT: Kim, Joonyul

TITLE OF INVENTION: Novel Carotenoid Hydroxylases for Use in Engineering Carotenoid

FILE OF INVENTION: Metabolism in Plants

FILE REFERENCE: MSU-08604

CURRENT APPLICATION NUMBER: US/10/751,235

CURRENT FILING DATE: 2004-01-02

NUMBER OF SEQ ID NOS: 74

SOFTWARE: Patent in version 3.2

SEQ ID NO 4

LENGTH: 539

TYPE: PRT

ORGANISM: Arabidopsis thaliana

US-10-751-235-4

Query Match 100.0%; Score 122; DB 5; Length 539;
Best Local Similarity 100.0%; Pred. No. 9.9e-11;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LVAEVSEFLFGSGFAIAGPLWTA 24
|||
DB 144 LVAEVSEFLFGSGFAIAGPLWTA 167
|||

RESULT 4

US-10-739-930-6317
Sequence 6317, Application US/10739930
Publication No. US20040216190A1
GENERAL INFORMATION:

APPLICANT: Kovalic, David K.

TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH

FILE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT

FILE REFERENCE: 38-21(53377)B

CURRENT APPLICATION NUMBER: US/10/739,930

CURRENT FILING DATE: 2003-12-18

NUMBER OF SEQ ID NOS: 11088

SEQ ID NO 6317

LENGTH: 560

TYPE: PRT

ORGANISM: Arabidopsis thaliana

FEATURE:

OTHER INFORMATION: Clone ID: ARATH-23APR03-C3001_1.p

US-10-739-930-6317

Query Match 100.0%; Score 122; DB 5; Length 560;
Best Local Similarity 100.0%; Pred. No. 1e-10;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LVAEVSEFLFGSGFAIAGPLWTA 24
|||
DB 165 LVAEVSEFLFGSGFAIAGPLWTA 188
|||

SUBSET OF STIC SEQUENCE SEARCH
FOR SEQ ID NO: 10

RESULT 5

US-10-751-235-21
Sequence 21, Application US/10751235
Publication No. US20050150002A1
GENERAL INFORMATION:

APPLICANT: Dellapenna, Dean

APPLICANT: Tian, Li

APPLICANT: Kim, Joonyul

TITLE OF INVENTION: Novel Carotenoid Hydroxylases for Use in Engineering Carotenoid

FILE OF INVENTION: Metabolism in Plants

FILE REFERENCE: MSU-08604

CURRENT APPLICATION NUMBER: US/10/751,235

CURRENT FILING DATE: 2004-01-02

NUMBER OF SEQ ID NOS: 74

SOFTWARE: Patent in version 3.2

SEQ ID NO 21

LENGTH: 208

TYPE: PRT

ORGANISM: Helianthus annuus

US-10-751-235-21

Query Match 93.4%; Score 114; DB 5; Length 208;
Best Local Similarity 95.8%; Pred. No. 6.6e-10;
Matches 23; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LVAEVSEFLFGSGFAIAGPLWTA 24
|||
DB 29 LVAEVSEFLFGSGFAIAGSLWTA 52
|||

RESULT 6

US-10-751-235-18
Sequence 18, Application US/10751235
Publication No. US20050150002A1
GENERAL INFORMATION:

APPLICANT: Dellapenna, Dean

APPLICANT: Tian, Li

APPLICANT: Kim, Joonyul

TITLE OF INVENTION: Novel Carotenoid Hydroxylases for Use in Engineering Carotenoid

FILE OF INVENTION: Metabolism in Plants

FILE REFERENCE: MSU-08604

CURRENT APPLICATION NUMBER: US/10/751,235

CURRENT FILING DATE: 2004-01-02

NUMBER OF SEQ ID NOS: 74

SOFTWARE: Patent in version 3.2

SEQ ID NO 18

LENGTH: 362

TYPE: PRT

ORGANISM: Triticum aestivum

US-10-751-235-18

Query Match 90.2%; Score 110; DB 5; Length 362;
Best Local Similarity 95.7%; Pred. No. 5.4e-09;
Matches 22; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LVAEVSEFLFGSGFAIAGPLWT 23
|||
DB 141 LVAEVSEFLFGSGFAIAGALWT 163
|||

RESULT 7

US-10-425-114-70634
Sequence 70634, Application US/10425114
Publication No. US20040034888A1
GENERAL INFORMATION:

APPLICANT: Liu, Jingdong

APPLICANT: Zhou, Yihua

APPLICANT: Kovalic, David K.

APPLICANT: Screen, Steven E

APPLICANT: Tabaska, Jack E

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
Plants and Uses Thereof for Plant Improvement

ORGANISM: Arabidopsis thaliana
US-10-751-235-4

Query Match 100.0%; Score 177; DB 5; Length 539;
Best Local Similarity 100.0%; Pred. No. 2.1e-14;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MESSLFPSSSSSYSLFTAKPTLLSPKPKFTFSIR 36
Db 1 MESSLFPSSSSSYSLFTAKPTLLSPKPKFTFSIR 36

RESULT 3
US-10-739-930-6317
Sequence 6317, Application US/10739930
Publication No. US20040216190A1
GENERAL INFORMATION:
APPLICANT: Kovacic, David K.
TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
FILE REFERENCE: 38-21(53377)B
CURRENT APPLICATION NUMBER: US/10/739,930
CURRENT FILING DATE: 2003-12-18
PRIOR FILING DATE: 2003-12-18
NUMBER OF SEQ ID NOS: 11088
SEQ ID NO 6317
LENGTH: 560
TYPE: PRT
ORGANISM: Arabidopsis thaliana
FEATURE:
OTHER INFORMATION: Clone ID: ARATH-23APR03-C3001_1.p
US-10-739-930-6317

Query Match 100.0%; Score 177; DB 5; Length 560;
Best Local Similarity 100.0%; Pred. No. 2.2e-14;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MESSLFPSSSSSYSLFTAKPTLLSPKPKFTFSIR 36
Db 22 MESSLFPSSSSSYSLFTAKPTLLSPKPKFTFSIR 57

RESULT 4
US-10-310-154-688
Sequence 688, Application US/10310154
Publication No. US20030233670A1
GENERAL INFORMATION:
APPLICANT: Edgerton, Michael D
APPLICANT: Chomet, Paul S.
APPLICANT: Adams, Thomas H
APPLICANT: Ruff, Thomas G.
APPLICANT: Agarwal, Ameeta K.
APPLICANT: Ahrens, Jeffrey E.
APPLICANT: Ball, James A.
APPLICANT: Banu, G.
APPLICANT: Bell, Erin
APPLICANT: Boddupalli, Raghava
APPLICANT: Deikman, Jill
APPLICANT: Deng, Molian
APPLICANT: Dong, Jinzhao
APPLICANT: Duff, Stephen M.
APPLICANT: Galligan, Meghan M.
APPLICANT: Hinchey, Brenda S.
APPLICANT: Huang, Shihshieh
APPLICANT: Johnson, G. Richard
APPLICANT: Jung, Vincent
APPLICANT: Kretzmer, Keith A
APPLICANT: Laccetti, Lucille B.
APPLICANT: Lai, Chao-Qiang
APPLICANT: Lee, Gary
APPLICANT: Lin, Jie-Yi
APPLICANT: Liu, Jingdong
APPLICANT: Lu, Bin
APPLICANT: Luethy, Michael M.

us-10-751-235-11.rapbm

APPLICANT: Lund, Adrian
APPLICANT: Madson, Linda L.
APPLICANT: Malloy, Kathleen A.
APPLICANT: McKiel, Christine L.
APPLICANT: Miller, Philip W.
APPLICANT: Padmavathi, Manchikanti
APPLICANT: Parnell, Laurence D.
APPLICANT: Start, William G.
APPLICANT: Tennesen, Dan
APPLICANT: Vidya, K.R.
APPLICANT: Wang, Haiyun
APPLICANT: Xin, Zhanqun
APPLICANT: Xu, Nanfei
APPLICANT: Yang, Chunzhi
APPLICANT: Zeng, Xiaoping
APPLICANT: Zhang, Qiang
APPLICANT: Zhao, Yajuan
APPLICANT: Zhou, Li
TITLE OF INVENTION: Gene Sequences and Uses Thereof in Plants
FILE REFERENCE: 38-15(52796)B
CURRENT APPLICATION NUMBER: US/10/310,154
CURRENT FILING DATE: 2002-12-04
PRIOR FILING DATE: 2001-12-04
NUMBER OF SEQ ID NOS: 736
SEQ ID NO 688
LENGTH: 260
TYPE: PRT
ORGANISM: Glycine max
US-10-310-154-688

Query Match 35.6%; Score 63; DB 4; Length 260;
Best Local Similarity 46.2%; Pred. No. 5.7;
Matches 12; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

Qy 10 SSSYSLFTAKPTLLSPKPKFTFSI 35
Db 4 SATSASLFSANPTPLFSPKPSLSLHL 29

RESULT 5
US-10-732-923-601
Sequence 601, Application US/10732923
Publication No. US20050108791A1
GENERAL INFORMATION:
APPLICANT: Edgerton, Michael D
TITLE OF INVENTION: TRANSGENIC PLANTS WITH IMPROVED PHENOTYPES
FILE REFERENCE: 38-15(52796)C
CURRENT APPLICATION NUMBER: US/10/732,923
CURRENT FILING DATE: 2003-12-10
PRIOR FILING DATE: 2002-12-04
NUMBER OF SEQ ID NOS: 24149
SEQ ID NO 601
LENGTH: 260
TYPE: PRT
ORGANISM: Glycine max
US-10-732-923-601

Query Match 35.6%; Score 63; DB 5; Length 260;
Best Local Similarity 46.2%; Pred. No. 5.7;
Matches 12; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

Qy 10 SSSYSLFTAKPTLLSPKPKFTFSI 35
Db 4 SATSASLFSANPTPLFSPKPSLSLHL 29

RESULT 6
US-10-732-923-21136
Sequence 21136, Application US/10732923
Publication No. US20050108791A1
GENERAL INFORMATION:

SUBSET OF STIC SEQUENCE SE ARCH
FOR SEQ IN NO=11